

U.S. DEPARTMENT OF ENERGY

# JOINT GENOME INSTITUTE

*sequencing the world of possibilities  
for energy and the environment*

The genome of an organism contains all of its genetic material or DNA, molecules made up of four bases known as A, C, T and G. When scientists sequence the genome, it's like taking apart a puzzle someone else completed to figure out what each section is made up of, and then re-assemble the information so that scientists can understand the big picture.

The U.S. Department of Energy Joint Genome Institute has been headquartered in Walnut Creek, California since 1999, sequencing plants, microbes and communities of microbes called metagenomes that are related to the DOE mission areas of bioenergy, the carbon cycle and biogeochemistry. A fifth of the sequencing projects done worldwide is done right here in Walnut Creek.

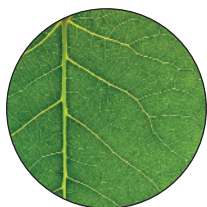
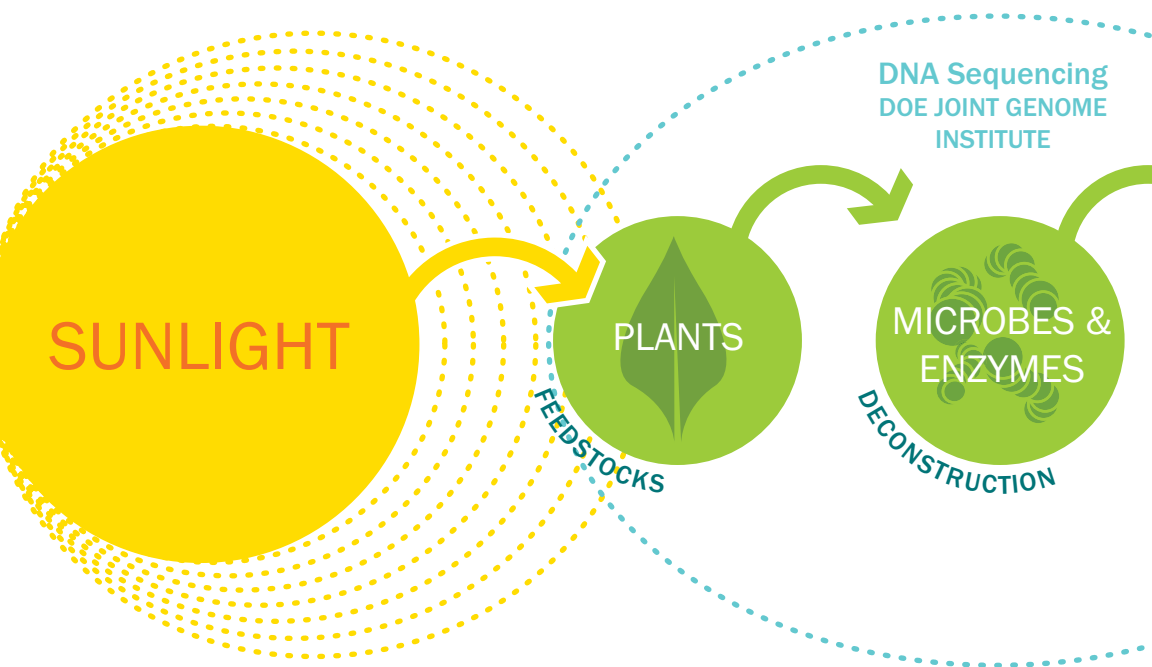


U.S. DEPARTMENT OF  
**ENERGY**  
Office of Science

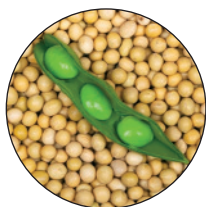
# BIOENERGY

The United States currently accounts for 21 percent of the world's total energy consumption, and imports 11 of the 20 million barrels of oil used daily.

Roughly two-thirds of the energy consumed by the United States is used by transportation and industry, so part of the DOE's commitment to energy security involves developing clean alternative energy sources that are commercially competitive with the oil industry. As a result, many DOE JGI sequencing projects are related to at least one of three aspects of biofuel production: developing plant feedstocks; using microbes to break down the cellulose in the plant cell walls; and fermenting the sugars into bioalcohols.



**Poplar genome:** DOE JGI sequenced and published the first tree genome — of the poplar tree — in 2006. Poplar is a fast growing tree that could be used as a biofuel feedstock, so researchers want to make it more resistant to both pests and drought, as well as increase its biomass.

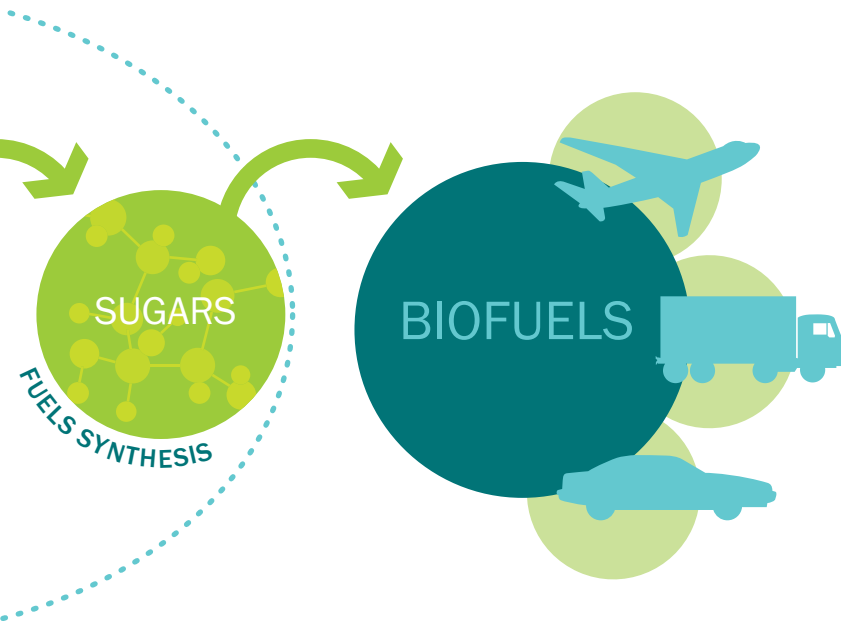


**Soybean genome:** Soybean is used for biodiesel, is second to corn as a biofuel feedstock in the United States while using a third less water. The sequencing information is used as a template to study other biofuel grass genomes that could also serve as feedstocks.

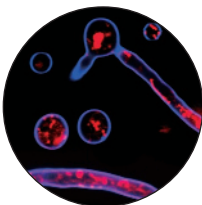


**Cow rumen metagenome:** In industry, the process currently used to break down plant biomass is expensive and involves harsh chemicals. Researchers want to identify the enzymes in the microbial community of the cow's first stomach that break down the grass it eats and apply them to industrial biofuel production.

*Photo courtesy of Peggy Greb, USDA*



**Termite hindgut metagenome:** The enzymes in the termite hindgut that make it possible for the tiny creature to destroy a house could be used to reduce the cost of breaking down plant biomass into simple sugars for cellulosic biofuel production. Another set of bacteria in the termite gut also produces hydrogen that combines with the simple sugars to produce the acetate the termites use for energy.

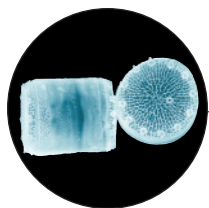
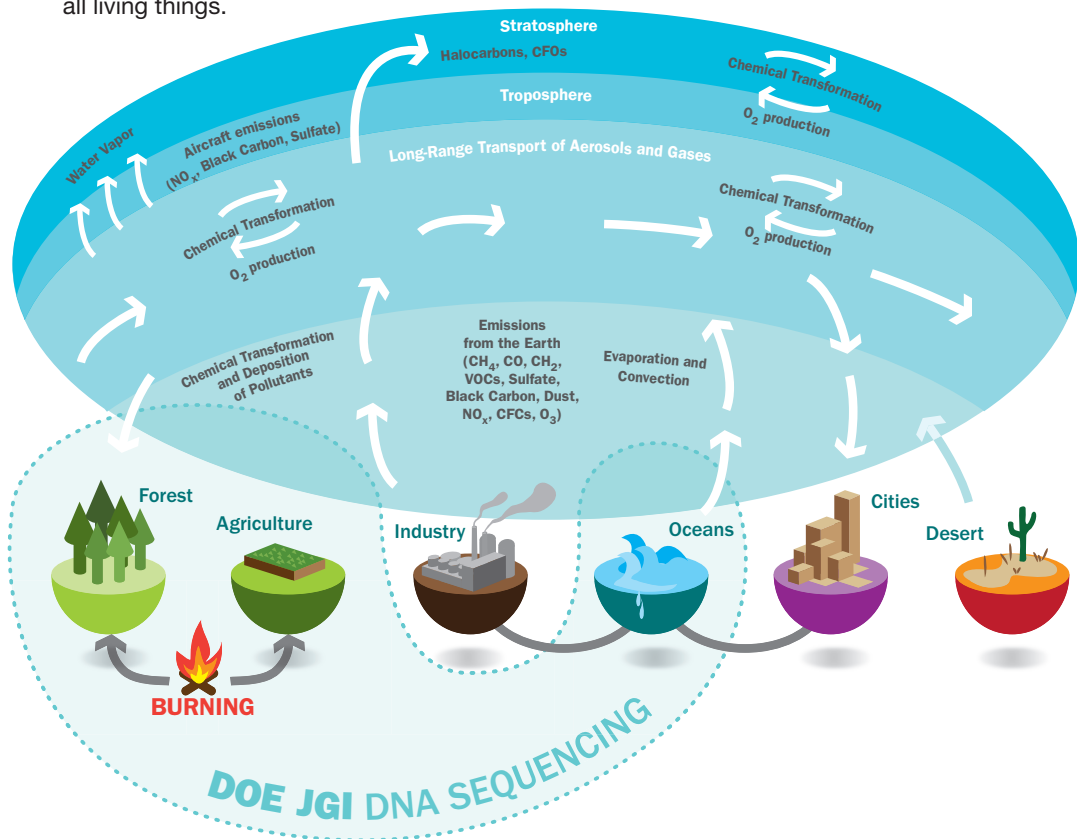


***Trichoderma reesei* genome:** The fungus produces enzymes that break down cellulose or cellulases. Industry already relies on fungal *T. reesei* strains that have increased ability to produce the cellulases, but further boosting enzyme production capabilities would benefit projects such as commercial biofuel production.

*Photo courtesy of Mari Valkonen, VTT Finland*

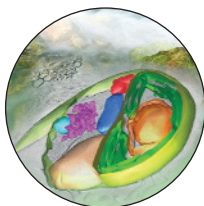
# CARBON CYCLE

Scientists are studying carbon sources and sinks on land, in the oceans and in the atmosphere to understand how changes, either natural or manmade, affect the planet and all living things.



***Thalassiosira pseudonana* genome:** Diatoms generate as much as 40 percent of the 50 billion to 55 billion tons of organic carbon produced each year in the sea and in the process use carbon dioxide and produce oxygen. Researchers study diatoms to understand how they can absorb carbon dioxide in amounts comparable to all the world's tropical rain forests combined.

Photo courtesy of Nils Kröger, Universität Regensburg



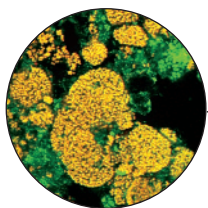
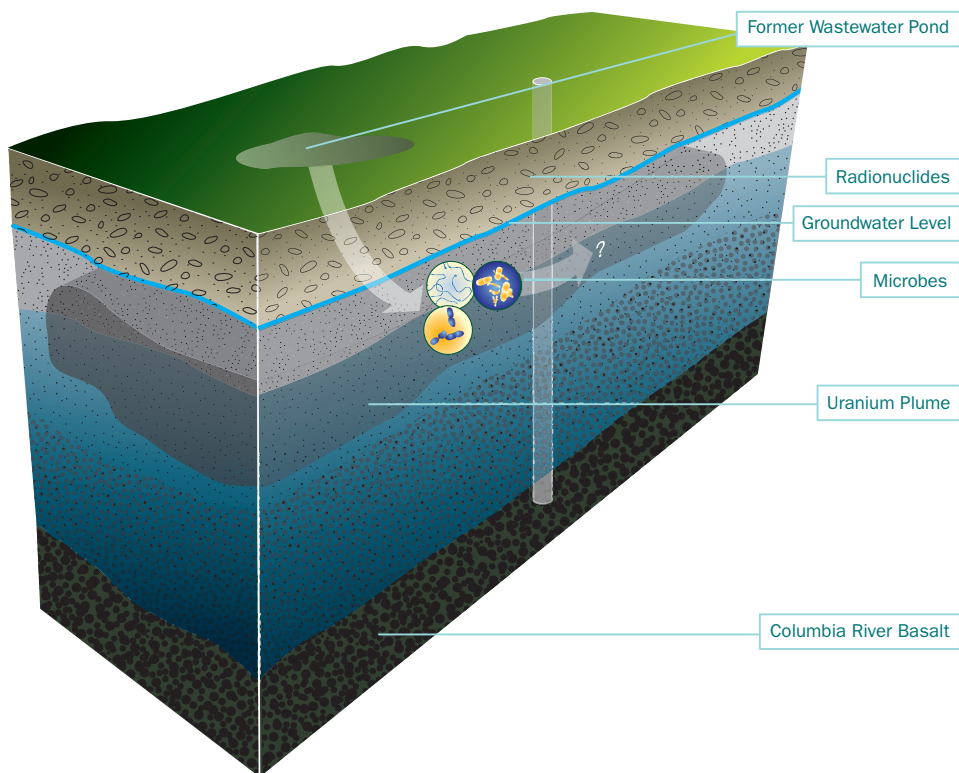
***Micromonas* genome:** Algae can capture atmospheric carbon and transport it from the ocean surface zones to the depths below. *Micromonas* are responsible for three-quarters of the net carbon production and strains from English Channel and South Pacific have shown they can adapt to environmental changes.

3D reconstruction by A.Z. Worden, T. Deerinck, M. Terada, J. Obiyashi and M. Ellisman, MBARI and NCMIR. Background image: Flavio Robles, Lawrence Berkeley National Laboratory



# BIOGEOCHEMISTRY

Biogeochemistry studies the global processes involved in making — and keeping — the Earth habitable. One aspect of this field is bioremediation, or using microbes to maintain and restore the natural environment.

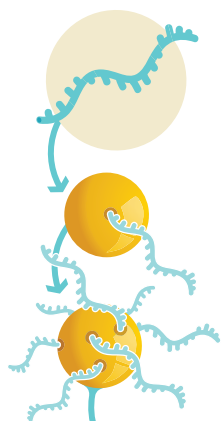


**Wastewater treatment:** DOE JGI researchers studied the microbial communities in activated wastewater sludge treatment process, focusing on the specialized bacteria that remove phosphorus from sewage to protect water systems from algal blooms. They were able to obtain a nearly complete genetic blueprint for a key player in this process, the bacterial species *Accumulibacter phosphatis* which had never before been isolated.

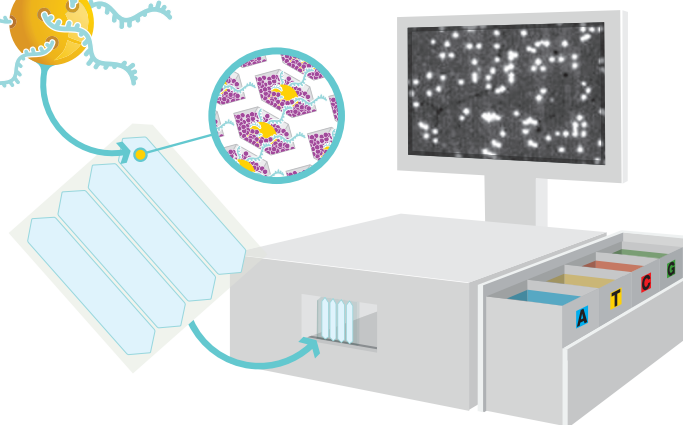
# SEQUENCING TECHNOLOGIES AT

A decade ago, the DOE JGI relied exclusively on the Sanger method to generate 20 million bases of sequence in a year, arranging and assembling DNA fragments by size. New sequencing technologies in use at the DOE JGI employ machines that can sequence more samples simultaneously and do it faster, generating a few billion bases of sequence in a week. Biotech start-ups are already hinting at machines that could sequence a few billion bases in less than an hour.

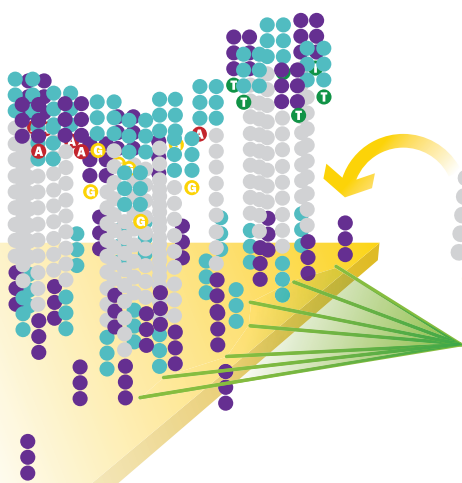
The new technologies determine the complement of the DNA fragment being sequenced first, and then researchers can work backwards to determine the original sequence.



**454:** DNA fragments are stuck to special beads and then amplified. Each bead then goes into a single well on a 3.6 million-well picotiter plate and is surrounded by enzymes. Each of the four nucleotides then flows over the wells, binding to the last base at the end of the DNA fragment attached to the bead. To detect the attached nucleotides, pyrosequencing produces light that is then captured by a camera, and the information is used to produce a DNA sequence that complements the original sequence the researchers were trying to puzzle out.



**Illumina:** Samples are loaded on to a tile on a 960-tile FlowCell for sequencing and undergo “bridge amplification” to generate single-stranded clusters of template DNA. All four nucleotides flow over the tiles, attaching themselves to bases on the DNA fragments. Lasers make the DNA clusters light up, allowing a camera to record the colored lights associated with each base. The information is used to produce the sequence of the DNA fragment, which is then aligned to a reference sequence to determine the full genome.



# THE DOE JGI

These developments lead to changes in the DOE JGI workforce as well as more focus on pre- and post-sequencing services. To truly sequence representative organisms, for example, researchers need to look at organisms that are currently difficult to cultivate in the lab.

Approximately 99.9 percent of the microbes in the biosphere are still unknown because most of them don't thrive once away from their natural habitats. DOE JGI scientists are developing techniques to collect samples from such organisms to sequence them.



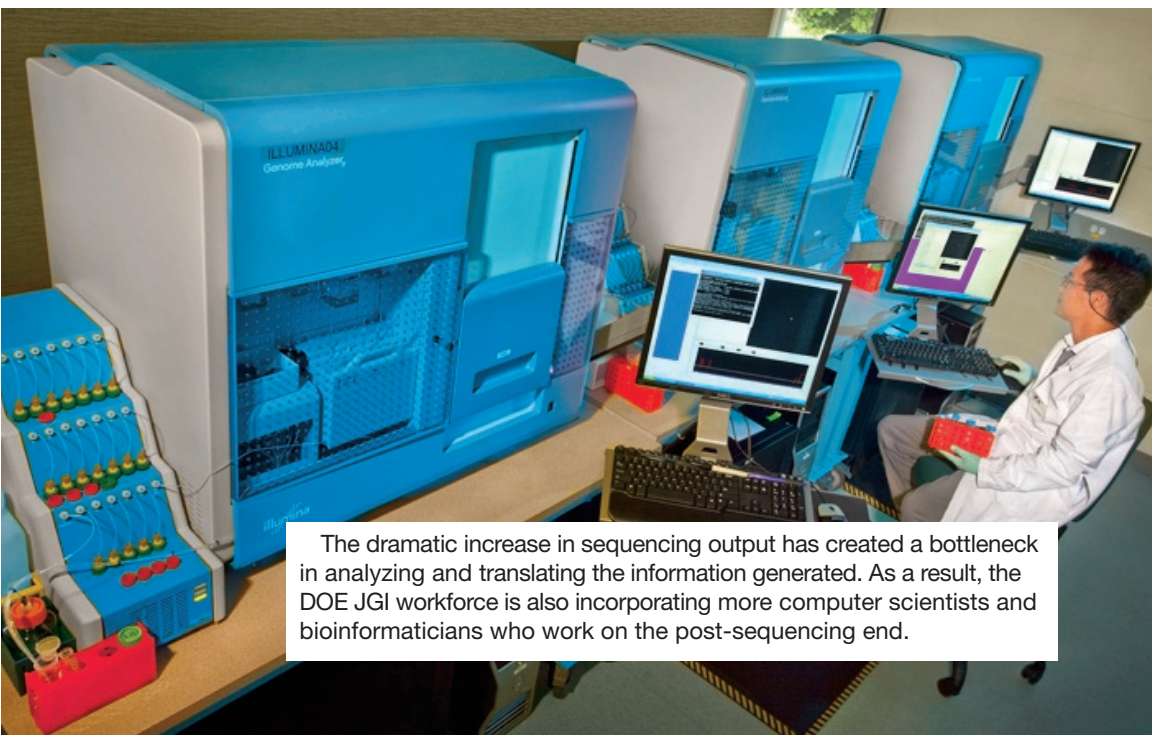
**Metagenomics:** DOE JGI looks at entire microbial communities to get a partial understanding of each microbe's structure and function when individual genome studies aren't possible. For example, researchers from Washington studied Lake Washington sediments to identify the biochemical pathways associated with nitrogen cycling and how microbes generate and consume the greenhouse gas and potential energy source methane.

*Photo courtesy of D. Kunkel and E. Latypova*



**Single cell sequencing:** taking the DNA from a single cell and generating enough copies to be able to sequence the hard-to-cultivate organism.

*Photo courtesy of Ramunas Stepanauskas, Bigelow Laboratory*



The dramatic increase in sequencing output has created a bottleneck in analyzing and translating the information generated. As a result, the DOE JGI workforce is also incorporating more computer scientists and bioinformaticians who work on the post-sequencing end.





Job openings at the DOE Joint Genome Institute (JGI) normally reflect the wide variety of projects undertaken by the Institute. Job categories include: Production, Research, Bioinformatics, Software Development, Operations, Finance and Procurement. Employees of the DOE JGI are usually hired through Lawrence Berkeley National Laboratory or Lawrence Livermore National Laboratory and are assigned to the DOE JGI through the various divisions within the Laboratories.

**Current opportunities may include:**

**Production Technicians  
Research / Sr. Research Associates  
Bioinformatics Data Analysts  
Software Developers  
Postdoctoral Researchers  
Research Scientists**

To view a complete list of current openings, including detailed job descriptions, visit [www.jgi.doe.gov](http://www.jgi.doe.gov) and click on the "Career" tab.

The DOE JGI is an equal opportunity employer with a commitment to workforce diversity.

**[www.jgi.doe.gov](http://www.jgi.doe.gov)**